DOCUMENT RESUME

ED 126 508 CS 202 839

AUTHOR Greenbaum, Sidney
TITLE Experiments on Judgements of Syntactic Acceptability and Frequency: Report 1.
PUB DATE 76
NOTE 21p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Descriptive Linguistics; English; Higher Education; *Language Patterns; Language Research; *Language Usage; Native Speakers; Research Methodology; *Sentence Structure; *Syntax

ABSTRACT

Overlapping samples of 191, 142, and 87 undergraduates rated pairs of sentences for the frequency and acceptability of the syntactic constructions represented to investigate the feasibility of using such ratings in the study of syntactic forms. The results indicated that subjects are consistent in their judgments of both the frequency and acceptability of the syntactic constructions tested, and that constructions rated high in frequency tend to be rated as more acceptable. (AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED126508

EXPERIMENTS ON JUDGEMENTS OF SYNTACTIC ACCEPTABILITY AND FREQUENCY: REPORT 1

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Sidney Greenbaum
Department of English
University of Wisconsin-Milwaukee
July 1976

INTRODUCTION

In devising syntactic description, linguists have generally relied for their primary data on the judgements of native speakers as to which constructions belong to the language. If he is a native speaker, the judgements are most frequently those of the linguist himself. The reliability of such intuitions has been assailed from a number of directions, particularly when the data are supplied by linguists (e.g. Bever 1970, 343-348; Levelt 1972; Labov 1972, 198f.; Greenbaum 1973), and there is now a body of evidence demonstrating that the methods of elicitation may affect judgements (e.g. Greenbaum 1973; Elliot et al. 1974; Greenbaum 1974). Partly as a result of concern over the reliability of acceptability intuitions, some transformational-generative linguists have been returning to the systematic collection of samples of language in actual use as the primary data for language description, but this trend has been mainly restricted to sociolinguistic studies of phonetic variables and of a few syntactic variables of very high frequency (cf. particularly Labov 1972, 207-216). It is likely that for the foreseeable future linguists will continue to use acceptability intuitions as an important source of data for syntactic description.

Corpus studies have long been used for discovering data to be included in a syntactic description. Linguists who emphasize a functional view

CS 208 539

of language description tend to be interested in relative frequency of language use. Such linguists have used corpus studies to provide data on relative frequencies of syntactic constructions in the language as a whole, or within a particular variety of the language, or between different varieties of the language (e.g. Leech 1966; Svartvik 1966). Surprisingly enough, linguists have not attempted to elicit judgements of frequency to supplement corpus data on frequency. Labov has denied that native speakers can report reliably on relative frequency, but his claim appears in the context of a discussion of variables that are strong markers of social values (Labov 1972, 266), and it may not be true for variables that have no such function.

The experiments reported on here were intended to investigate (a) whether it is feasible to use frequency judgements by native speakers for the syntactic description of English, and (b) whether there is a relationship between frequency and acceptability judgements. The sentences included in the experiments constituted several sets of syntactic options in English. It was hoped that the results would indicate the conditions that influence changes in perceived frequencies.

SUBJECTS AND PROCEDURE

Two sets of experiments were conducted, each set consisting of two experiments eliciting judgements that were to be correlated:

1a - Scale Frequency judgements

1b - Proportional Frequency judgements

2a - Scale Frequency judgements

2b - Acceptability judgements

This report will be concerned only with the Scale Frequency and Acceptability judgements in experiments 1a, 2a and 2b.

For an analysis of the Frequency judgements alone, we conflate the results of the 1a and 2a experiments, since there was no selection involved in the individuals appearing in the two experiments and the procedures were the same on both occasions. Three sets of results are therefore analyzed in this report:

(1) 1a and 2a - Frequency judgements

(2) 2b - Acceptability judgements

(3) 2a and 2b - Frequency judgements correlated with Acceptability judgements

The experiments were given in May-April 1974 to students attending large History lecture courses at the freshman level at the University of Wisconsin-Milwaukee. With the permission of the instructors, the investigator conducted the experiment at the beginning of lecture periods.[1] The subjects, who participated voluntarily, were told that the experiments were concerned with how language works. The a and b experiments in the same set were given at the same lecture course at least a week apart.

The subjects in the experiments were nonlinguists, predominantly freshmen and sophomores and predominantly from the State of Wisconsin. The majority were males between the ages of 18 and 21. The numbers for each set of results were: 1a and 2a - 191; 2b - 142; 2a and 2b - 87.[2]

The sentences presented to the subjects were identical for both the Frequency and Acceptability experiments. They are listed in Appendix 1.

The sentences were given in pairs in a stapled booklet. Each of the alternate orders was given to half of the subjects. The order of the pairs was randomized afresh for each booklet.

The subjects were untimed, though they were encouraged to work quickly. Appendix II gives the instructions for the two types of experiments, which were presented at the beginning of each booklet. As can be seen from the instructions, the subjects were asked to judge each sentence in a pair on a five-point scale. The extremes for the Frequency experiment were marked very rare and very frequent, while those for the Acceptability experiment were marked completely unacceptable and perfectly OK.

MATERIAL

The 50 pairs of sentences are grouped linguistically in Appendix 1. Subjects judged sentences in pairs. Sentences within a pair have the same lexical content but differ syntactically. It was hoped that judgements would focus on the syntactic variation within the pair. The 25 numbered sets each have two pairs of sentences: an a-b pair and a c-d pair. The c-d pairs repeat the syntactic variation in the corresponding a-b pairs but with a different lexical content; for example, the active/passive contrast in 1a-1b (1a: Marvin saw Susan - 1b: Susan was seen by Marvin) recurs with lexical variation in 1c-1d (1c: Bruce called Jane - 1d: Jane was called by Bruce). The c-d pairs provide some measure of control over the extent to which the judgements of frequency and acceptability are influenced by the lexical content of the pair of sentences. In the experiments the c-d pairs appeared in the second half of the booklet, in the same sequence as the randomized a-b pairs.

The numbered sets are here grouped into eight sections that are labeled to indicate the syntactic variation being investigated. Thus, the first section is concerned with the option between active and passive in English. In set 1 the sentences are declarative, in set 2 they are interrogative, and in set 3 they are declarative but also contain a modal auxiliary. At issue is whether these syntactic differences affect judgements on the acceptability and frequency of the sentences. For example, is the perceived frequency of the active/passive contrast different for questions and statements? Report 2 contains a linguistic analysis of this type of data in the results.

RESULTS

Intra-subject Consistency

Subjects were required to indicate their judgements by putting a check-mark on each line, in one out of five boxes. A subject can be said to be consistent in his judgements if he puts his check-mark in an identical box on two occasions ('Direct Hits'). A weaker acknowledgement of consistency allows for adjacent boxes as well; that is to say, the subject is considered consistent not only if his check-mark is in identical boxes on both occasions but also if the check-mark in the subsequent experiment is in a box immediately higher or immediately lower than it was in the previous experiment ('Direct Hits ± 1').

Table 1 summarizes the results of intra-subject consistency for three sets of data:

(1) frequency judgements of sentences that have different lexical content but are considered to be identical in the syntactic feature under investigation, i.e. each a sentence and its corresponding c sentence, and each b sentence and its corresponding d sentence;

(2) acceptability judgements of the same sets of sentences as in (1);

(3) the frequency judgement of each sentence and the acceptability judgement for the same sentence.

The first column in each set of results in Table 1 gives the number of related sentences for which a given percentage range of subjects achieved Direct Hits; for example, in the first column for Frequency judgements, there were three pairs of sentences where 70-74% of the subjects had Direct Hits. The second column in each set shows the same for Direct Hits ±1.

Absolute consistency is perhaps not to be expected for judgements where either the sentences varied (sets 1 and 2) or the type of judgement varied (set 3), but it might be noted that in 27 of the 50 pairs of sentences 50% or more of the subjects gave identical Acceptability judgements for the lexically-varied sentences. If we use the weaker measure of consistency and allow a one-position difference between two judgements, the results are impressive. In 46 of the 50 pairs of sentences in the Frequency experiment 70% or more of the subjects achieved Direct Hits ±1, while in the Acceptability experiment there were 47 pairs within that percentage. For set 3, in 88 of the 100 sentences 65% or more achieved Direct Hits ±1 when they judged both for Frequency and for Acceptability.

As general measures for the data, we can devise an intra-subject consistency index: we add percentages of subjects achieving Hits and divide that number by the number of tests. The intra-subject consistency indices for the three sets are:

	Frequency	Acceptability	Frequency v. Acceptability
Direct Hits	44.3%	51.5%	37.2%
Direct Hits ±1	79.3%	79.9%	74.6%

TABLE 1

INTRA-SUBJECT CONSISTENCY: SCORES

% OF SUBJECTS	FREQUENCY (1a + 2a) Lexically-varied Sentences (50 pairs)		ACCEPTABILITY (2b) Lexically-varied Sentences (50 pairs)		FREQUENCY (2a) v. ACCEPTABILITY (2b) (100 pairs)	
	Direct Hits	Direct Hits ± 1	Direct Hits	Direct Hits ± 1	Direct Hits	Direct Hits ± 1
95–100		2				4
90–94		2		4		5
85–89		8		9		2
80–84		9		11		13
75–79		14	2	11	1	19
70–74	3	11	1	12		25
65–69		3	3	3	3	20
60–64	2	1	5		2	9
55–59	3		8		1	3
50–54	4		8		2	
45–49	7		8		8	
40–44	9		8		12	
35–39	12		3		23	
30–34	10		4		24	
25–29					20	
20–24					4	

The results indicate that subjects can generally be consistent in their syntactic judgements of both Frequency and Acceptability of sentences that are designed to be similar syntactically though differing in lexical content, if the measure of consistency allows for a one-position difference on the five-place scale used on the two occasions. The results also suggest that there is an association between frequency and acceptability judgements.

For the set 3 data we can also consider intra-subject consistency in the direction of judgements within pairs of sentences. That is to say, if sentence a (e.g. 1a: Marvin saw Susan) is judged to be more frequent than sentence b (e.g. 1b: Susan was seen by Marvin), we can ask whether a is judged more acceptable than b. Two consistency measures were again devised. A subject is consistent if on both occasions he marked the same sentence as higher than the other in the pair or if he marked both as equal on the scale. He is of course totally inconsistent if he judged the frequency of a as higher than b but judged the acceptability of a as lower than b. A weaker measure of inconsistency allows the subject to mark the same position for the two sentences on one occasion and different positions on the other occasion ('partial consistency'); for example, he judged a and b as equally frequent, but a as more acceptable than b. The three columns in Table 2 give the percentage range, the number of sentences where there is total consistency in direction of judgements, and the number of sentences where there is either total consistency or partial consistency.

TABLE 2

INTRA-SUBJECT CONSISTENCY:
DIRECTION OF FREQUENCY AND ACCEPTABILITY JUDGEMENTS

% OF SUBJECTS	TOTAL AGREEMENT IN DIRECTION FOR 2 PAIRS	TOTAL AGREEMENT + CASES OF ONE PAIR WITH EQUAL SCORES
95-100		4
90-94		4
85-89		8
80-84	3	10
75-79	3	17
70-74	2	7
65-69	3	
60-64	5	
55-59	5	
50-54	11	
45-49	5	
40-44	8	
35-39	5	

As can be seen from the table, in 32 of the pairs over 50% of the subjects marked the sentences in the same direction for both judgements, while in all 50 pairs over 70% of the subjects were either totally or partially consistent. Total inconsistency ranged for individual pairs of sentences from 7.6% of the subjects to 28.6%. The intra-subject consistency index for total consistency is 55.1%, while for combined total and partial consistency it is 81.7%; the index for total inconsistency is 18.2%. The analysis provides support for the hypothesis that frequency and acceptability ratings for a given pair of sentences tend to go in the same direction.

A similar indication is conveyed by an analysis of the mean scores, listed in Table 3. The mean scores for the two types of judgement on individual sentences differed less than the value of one position in the five-place scale for any of the 100 sentences. The greatest difference was .94, but in 67 sentences it was less than .5, and in as many as 22 sentences it was less than .1. The difference was predominantly in the direction of the Acceptability mean score being higher than the Frequency mean score: 86 out of 100 sets of judgements. Of the 14 sets where the reverse occurred - Frequency mean score higher than Acceptability mean score (indicated by an arrow on the Table) - the mean difference was less than .21. The mean scores point to a narrower range for Acceptability judgements than for Frequency judgements. In both cases, the highest mean score is 4.69; but the lowest mean score for Acceptability is 2.5, whereas for Frequency there are 16 sentences with mean scores below 2.5 and the lowest mean is 1.57.

TABLE 3

MEAN SCORES FOR ACCEPTABILITY AND FREQUENCY JUDGEMENTS:
SUBJECTS IN BOTH EXPERIMENTS

a-b pairs						c-d pairs					
	Acc.	Freq.		Acc.	Freq.		Acc.	Freq.		Acc.	Freq.
1a	4.57 →	4.60	14a	4.19	3.99	1c	4.60 →	4.66	14c	4.09	4.01
1b	2.67	1.74	14b	3.00	2.31	1d	2.65	1.93	14d	2.66	2.35
2a	4.52 →	4.69	15a	2.95	2.40	2c	4.69	4.67	15c	3.19	2.56
2b	2.59	1.67	15b	4.20	4.08	2d	2.53	1.89	15d	4.16 →	4.24
3a	4.51 →	4.57	16a	4.09	3.85	3c	4.47 →	4.52	16c	3.94	3.67
3b	2.50	1.57	16b	3.38	3.05	3d	2.64	2.02	16d	4.02	3.40
4a	3.84	3.10	17a	4.28	4.08	4c	3.87	3.23	17c	3.71	3.13
4b	3.67	3.57	17b	3.38	2.69	4d	3.81	3.64	17d	3.94	3.64
5a	3.66	2.95	18a	2.66	1.94	5c	3.78	3.13	18c	2.94	2.17
5b	3.63	3.42	18b	4.45	4.45	5d	3.66	3.59	18d	4.26	4.02
6a	3.93	3.54	19a	4.17	4.10	6c	4.28	3.98	19c	3.95 →	4.16
6b	3.27	3.23	19b	4.05	3.45	6d	3.22	3.01	19d	3.97	3.16
7a	4.31	3.74	20a	4.02 →	4.09	7c	4.34	4.18	20c	3.87 →	3.94
7b	3.00	2.90	20b	3.98	3.43	7d	2.62	2.43	20d	3.92	3.30
8a	3.09 →	3.09	21a	4.23	4.18	8c	2.82 →	2.92	21c	4.20	4.17
8b	4.05	3.64	21b	3.57	3.21	8d	4.26	3.74	21d	3.72	3.21
9a	4.19	3.55	22a	3.33	3.05	9c	4.13	3.86	22c	3.33	2.43
9b	3.31	3.23	22b	3.69	3.51	9d	3.07	2.63	22d	4.04 →	4.09
10a	2.84 →	2.98	23a	4.28	3.90	10c	2.80 →	2.84	23c	4.28	3.82
10b	4.16	3.63	23b	2.88	2.45	10d	4.36	3.93	23d	3.00	2.48
11a	4.08	3.57	24a	4.24	4.12	11c	4.27	3.94	24c	3.89	3.66
11b	3.02	2.86	24b	3.63	2.99	11d	3.32	2.75	24d	3.69	3.29
12a	4.15	4.07	25a	3.58	3.16	12c	3.99	3.69	25c	3.36	3.16
12b	3.19	2.48	25b	3.68	3.41	12d	3.60	3.31	25d	4.03	3.67
13a	3.66	3.26				13c	3.36	2.95			
13b	3.80	3.60				13d	3.87	3.63			

An analysis of variance was used to estimate the reliability of the measurements (Winer 1971, 283-289). It was found to be more manageable for this purpose to separate the a and c sentences from the b and d sentences. Two reliability figures are given for eight sets of data: r = the reliability of the average of the given number of judges (i.e. the confidence that we can have that replication with a similar population of the same size would yield the same results) and r1 = the reliability of a single judge.

TABLE 4

RELIABILITY OF MEASUREMENTS

			R	R1
ALL SUBJECTS	Frequency (P1a & P2a)	a + c	.94462	.28324
		b + d	.96385	.27333
	Acceptability (P2b)	a + c	.96172	.19309
		b + d	.95844	.18293
SUBJECTS IN BOTH EXPERIMENTS	Frequency (P2a)	a + c	.96618	.27856
		b + d	.96512	.27214
	Acceptability (P2b)	a + c	.94833	.20772
		b + d	.94774	.20813

The reliability of a single subject is predictably low, since we know that there is considerable variation between individuals in acceptability judgements and we might expect the same variation for frequency judgements. The reliability of the group, however, is comfortingly high: we can cite the results with a great deal of certainty.

NOTES

[1] I am grateful to the instructors who gave me the opportunity to conduct the experiments: Professors Frank Cassell, Martin Schmidt, John Schroeder, and Roland Stromberg, all of the Department of History at the University of Wisconsin-Milwaukee. I am heavily indebted to Professor Robert Remstad of the Department of Educational Psychology for frequent advice on statistics and to Paul Kreuler of the Social Science Research Facility of the College of Letters and Science for an immense amount of computational work. I am also indebted to a number of students who helped in the administration of the tests or in the scoring of the results. This paper is a slightly revised version of the paper mimeographed in February 1975.

[2] The subjects were asked not to give their names, but they were required to devise an identification number to match the performance of individual subjects across experiments in the same set. They were also asked to provide certain biographical information: age, sex, year of studies, major, and states in which they had lived for at least one year. Most of the biographical information was elicited for the a experiments only, in the belief that approximately the same number of students would attend the lectures on both occasions. Unfortunately, a large proportion of the subjects were present on only one of the two occasions; as a result, the only biographical information available for those who attended just the b experiments is their age, since that was included in the identification number that the subjects devised.

Ages: 18-21 1a & 2a - 161 (84.3%); 2a & 2b - 81 (93.1%); 2b - 124 (87.3%)

Sex: Males 1a &2a - 139 (72.8%); 2a & 2b - 58 (66.7%)

Females 1a & 2a - 51 (26.7%); 2a & 2b - 29 (33.3%)

Freshmen or Sophomores: 1a & 2a - 142 (73%); 2a & 2b - 75 (86.2%)

Majors: Only a small number declared their major, since most had not yet begun to specialize: 1a & 2a - 45 (23.7%); 2a & 2b - 10 (11.5%); the majority were in the Social Sciences.

State: Wisconsin 1a & 2a - 165 (86.4%); 2a & 2b - 80 (92%)

Other Mid-west States 1a & 2a - 10 (5%), 2a & 2b - 5 (5.8%)

REFERENCES

Bever, T.G. (1970), 'The Cognitive Basis for Linguistic Structures', in Cognition and the Development of Language, ed. J.R. Hayes. The Hague.

Elliot, D., S. Legum, and S.A. Thompson (1974), 'Considerations in the Analysis of Syntactic Variation' (mimeo). SWRL Educational Research and Development. Los Alamitos, California.

Greenbaum, S. (1973), 'Informant Elicitation of Data on Syntactic Variation', Lingua 31.201-212.

Greenbaum, S. (1976), 'Contextual Influence on Acceptability Judgements', International Journal of Psycholinguistics.

Labov, W. (1972), Sociolinguistic Patterns. Philadelphia.

Leech, G.N. (1966), English in Advertising: A Linguistic Study of Advertising in Great Britain. London.

Levelt, W.J.M. (1972), 'Some psychological aspects of linguistic data', Linguistische Berichte 17.18-30.

Svartvik, J. (1966), On Voice in the English Verb. The Hague.

Winer, B.J. (1962), Statistical Principles in Experimental Design, second ed. New York.

APPENDIX 1

List of Sentences According to Linguistic Categories

Passive/Active

1a Marvin saw Susan.
b Susan was seen by Marvin.
c Bruce called Jane.
d Jane was called by Bruce.

2a Did Marvin see Susan?
b Was Susan seen by Marvin?
c Did Bruce call Jane?
d Was Jane called by Bruce?

3a Peter can watch the game.
b The game can be watched by Peter.
c Carol can read the book.
d The book can be read by Carol.

Modal Negation

4a I may be ready before you.
b I might be ready before you.
c She may be their teacher.
d She might be their teacher.

5a I may not be ready before you.
b I might not be ready before you.
c She may not be their teacher.
d She might not be their teacher.

Perfect/Past

6a He has spoken to me several times since he came here.
b He spoke to me several times since he came here.
c She has written to us a number of times since she moved to London.
d She wrote to us a number of times since she moved to London.

7a He has been sick ever since he left school.
b He was sick ever since he left school.
c She has been happy ever since she married Bob.
d She was happy ever since she married Bob.

Indicative/Subjunctive/SHOULD

8a We recommend that he pays full tuition.
b We recommend that he pay full tuition.
c We urge that he gives his reasons.
d We urge that he give his reasons.

9a We recommend that he pay full tuition.
b We recommend that he should pay full tuition.
c We urge that he give his reasons.
d We urge that he should give his reasons.

10a They demand that she reports all the facts.
b They demand that she report all the facts.
c They insist that she leaves the country.
d They insist that she leave the country.

11a They demand that she report all the facts.
b They demand that she should report all the facts.
c They insist that she leave the country.
d They insist that she should leave the country.

Relative Positions of Adverbials

12a We were waiting for three hours on Monday.
b We were waiting on Monday for three hours.
c They played tennis for twenty minutes yesterday evening.
d They played tennis yesterday evening for twenty minutes.

13a She goes swimming every afternoon for a few minutes.
b She goes swimming for a few minutes every afternoon.
c He has been staying with us every year for a week.
d He has been staying with us for a week every year.

14a I go there about three or four times during the vacation.
b I go there during the vacation about three or four times.
c Our electricity was cut off several times two months ago.
d Our electricity was cut off two months ago several times.

15a He was working before dinner in the garage.
b He was working in the garage before dinner.
c I eat my lunch on Wednesdays at the office.
d I eat my lunch at the office on Wednesdays.

16a He is lecturing in the next building after lunch.
b He is lecturing after lunch in the next building.
c They sell beer at the student cafeteria every day.
d They sell beer every day at the student cafeteria.

17a I spoke to him about three or four times last week.
b I spoke to him last week about three or four times.
c They came to our meetings on several occasions last year.
d They came to our meetings last year on several occasions.

Negative Contractions

18a They've not had that brand for years.
b They haven't had that brand for years.
c They've not heard what happened.
d They haven't heard what happened.

19a He's not here.
b He isn't here.
c She's not outside.
d She isn't outside.

20a Pete's not here.
b Pete isn't here.
c Kate's not outside.
d Kate isn't outside.

21a We're not going.
b We aren't going.
c We're not playing.
d We aren't playing.

Ellipsis

22a Some students have complained to the Chairman and they have written to the Dean.
b Some students have complained to the Chairman and written to the Dean.
c Our children will see the movie and they will visit the zoo.
d Our children will see the movie and visit the zoo.

23a Some students have complained to the Chairman and others have written to the Dean.
b Some students have complained to the Chairman and others written to the Dean.
c Our children will see the movie and the other children will visit the zoo.
d Our children will see the movie and the other children visit the zoo.

Indirect Object

24a Tom gave the boy a dime.
b Tome gave a dime to the boy.
c John lent the girl a book.
d John lent a book to the girl.

25a Tom gave a boy the dime.
b Tom gave the dime to a boy.
c John lent a girl the book.
d John lent the book to a girl.

APPENDIX 2

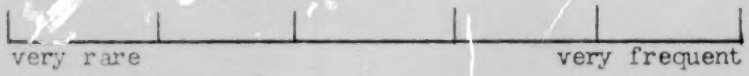
Instructions for Scale Frequency Experiment

The purpose of this study is to measure the awareness people have of the frequency of grammatical forms and constructions by having them judge their frequency on a scale. In making your judgements, please think of the overall frequency in the English language--not merely of your own use.

On each page of this booklet you will find a pair of sentences, like the following:

(a) John stands in the corner

(b) John is standing in the corner

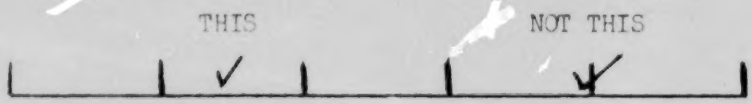


All you have to do is judge the frequency in the English Language for each sentence by putting a check-mark (✓) on each line, either in the leftmost box ('very rare') or the rightmost box ('very frequent') or one of those between.

We are interested in the constructions rather than in the individual words, so the frequency of the particular vocabulary is not relevant. For example, from that point of view it presumably makes no difference to your response if the sentences above contained Your brother instead of John. To help you restrict your judgement to grammatical frequency, we have put on each page a pair of sentences that differ only grammatically, for example in the forms of the words or in the order of the words.

IMPORTANT:

(1) Place your check-marks in the middle of spaces, not on the boundaries:



(2) Be sure you check every scale--do not leave any out!

Some of the sentences are similar. Do not turn back to previous pages, and do not try to remember how you checked similar sentences earlier in this experiment. When you are told to start, work through each page as quickly as you can. It is your first impressions that we want. On the other hand, please do not be careless, because we want your true impressions.

Instructions for Acceptability Experiment

The purpose of this experiment is to measure the attitudes that various people have towards grammatical forms and constructions by having them judge their acceptability on a scale. In making your judgements, please give your own feelings—not what you think others have said or might say.

On each page of this booklet you will find a pair of sentences, like the following:

(a) John stands in the corner

completely unacceptable — perfectly OK

(b) John is standing in the corner

completely unacceptable — perfectly OK

All you have to do is judge the acceptability of each sentence by putting a check-mark (✓) on each line, either in the leftmost box ('completely unacceptable') or the rightmost box ('perfectly okay') or one of those between.

IMPORTANT:

(1) Place your check-marks in the middle of spaces, not on the boundaries:

(2) Be sure you check every scale—do not leave any out!

Some of the pairs of sentences are similar. Do not turn back to previous pages, and do not try to remember how you checked similar sentences earlier in the experiment. When you are told to start, work through each page as quickly as you can. It is your first impressions that we want. On the other hand, please do not be careless, because we want your true impressions.